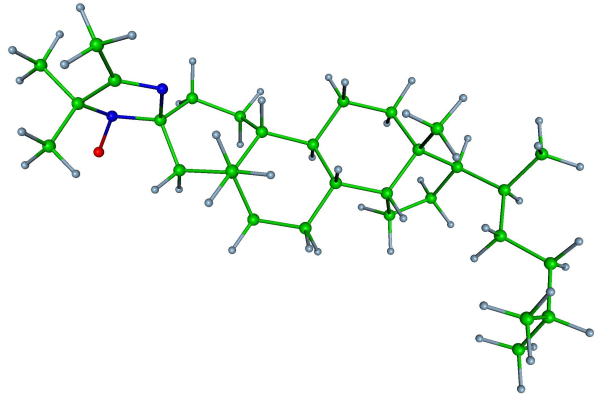
Recent Progress in Conformation Dynamics

Susanna Röblitz

Zuse Institute Berlin



DFG Research Center

Matheon

Overview

1. From Molecular Dynamics to Conformation Dynamics
2. Discretization of Position Space
3. Perron Cluster Analysis
4. Error Analysis
5. Softwarecode ZIBgridfree
6. Examples

Molecular Dynamics

Modelling of molecules in classical MD:

$$H(q,p) = \frac{1}{2} p^\top M^{-1} p + V(q)$$

$$V = V_{\text{bond}} + V_{\text{angle}} + V_{\text{torsion}} + V_{\text{Coulomb}} + V_{\text{VdW}}$$

$q \in \mathbb{R}^{3s}$: positions of all atoms, $p \in \mathbb{R}^{3s}$: momenta

Corresponding equations of motion:

$$\dot{q} = M^{-1} p, \qquad \dot{p} = -\nabla V$$

Ensemble averages

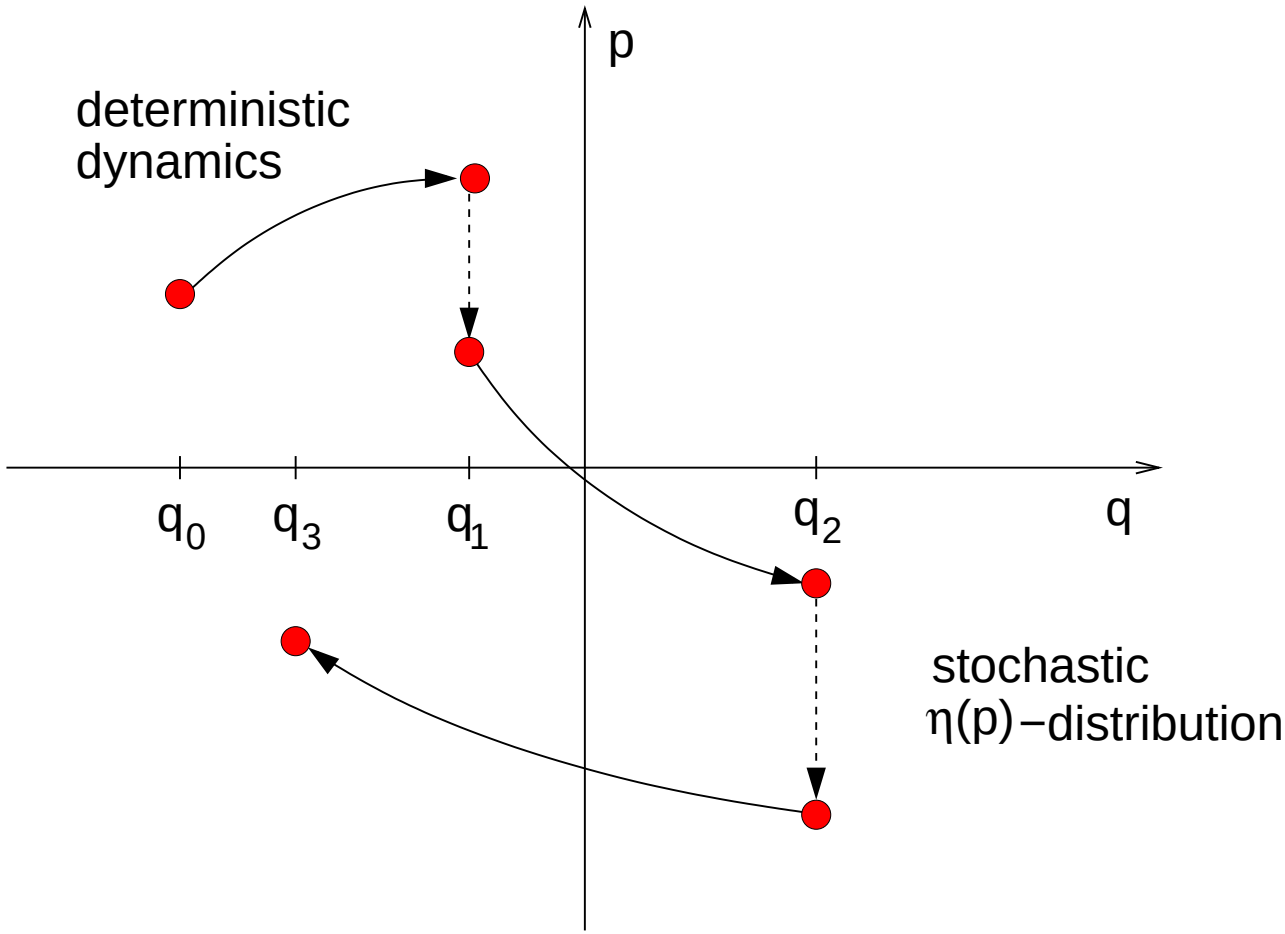
Invariant density in the canonical ensemble (Boltzmann):

$$\mu(q,p) = \underbrace{\frac{1}{Z_p}\exp\left(-\frac{\beta}{2}p^\top M^{-1}p\right)}_{=\eta(p)}\underbrace{\frac{1}{Z_q}\exp(-\beta V(q))}_{=\pi(q)}, \quad \beta = 1/k_B\mathcal{T}$$

Ensemble averages:

$$\langle A\rangle \equiv \int_\Omega A(q)\pi(q)\,dq \approx \frac{1}{M}\sum_{k=1}^{M} A(q_k), \quad q_k \sim \pi(q)$$

Hybrid Monte Carlo (HMC)



[Duane, Kennedy, Pendleton,Roweth (1987)]

Propagators

$$(q(\tau), p(\tau)) = \Psi^\tau(q(0), p(0))$$

required properties of the integrator Ψ^τ:

- time reversibility
- volume preservation

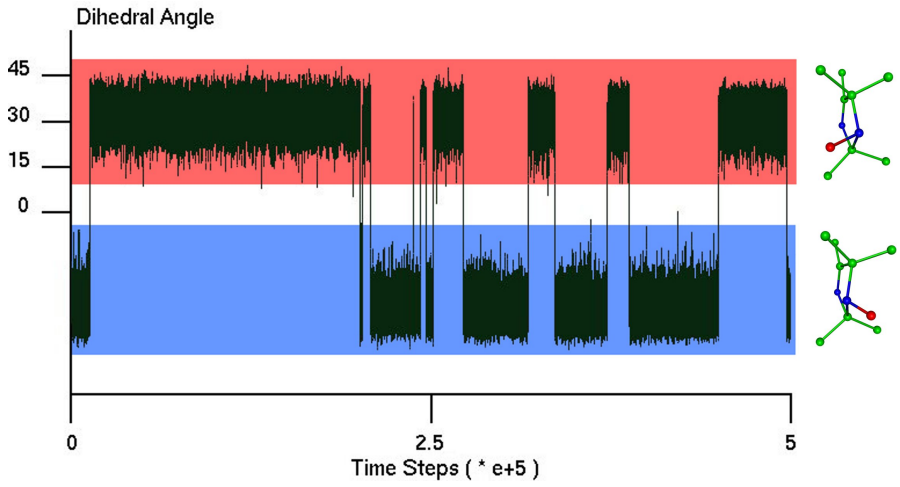
Example: Velocity Verlet integrator

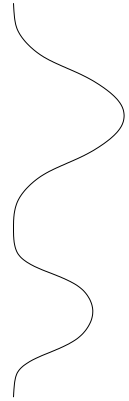
many possibilities for efficient and smart sampling

- full-flexible modeling of crucial parts of the system
- restricted or restrained modeling of less important parts

$\rightarrow$ gain in efficiency ($\mathcal{O}(s)$ instead of $\mathcal{O}(s^2)$) and running time

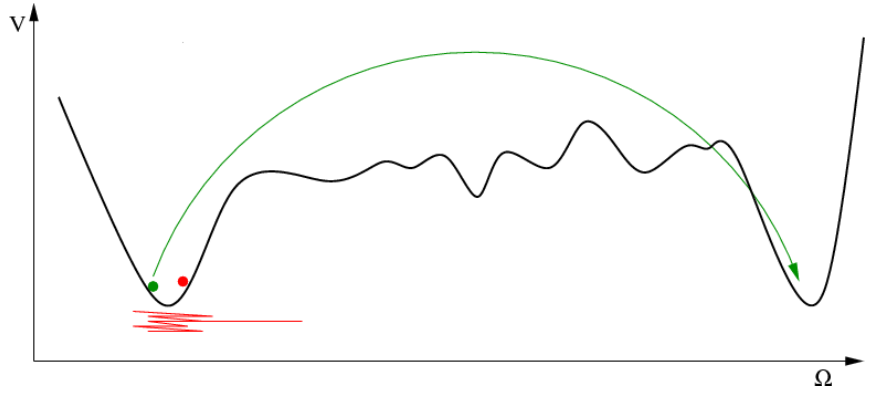
Metastable dynamical systems





stationary density?

The trajectory gets trapped in valleys of the potential energy surface.



Conceptual change

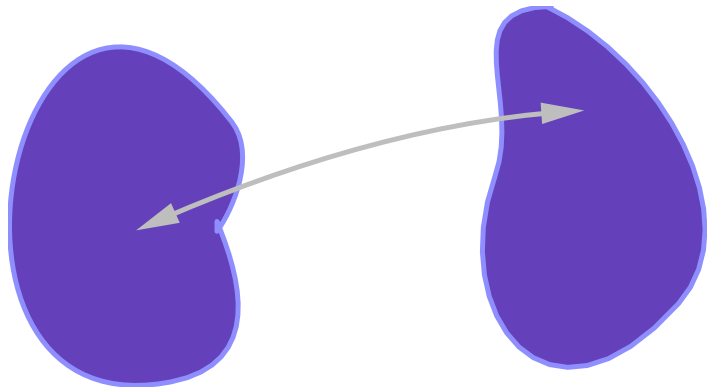
molecular dynamics

point concept:
trajectory simulation

deterministic

conformation dynamics

set concept:
metastable conformations



stochastic

[Deuflhard, Schütte (1997)]

Conformations as almost invariant sets

Transfer operator:

$$T^\tau u(q) = \int_{\mathbb{R}^d} u(\Pi_q \Psi^{-\tau}(q,p))\eta(p)\,dp$$

[Schütte, Fischer, Huisinga, Deuflhard 1998]

- Probability to be within A
  $w(A) = \int_A \pi(q)dq = \langle \chi_A, \chi_A \rangle_\pi$
- Probability to move from A $\longrightarrow$ B during time τ
  $w(A,B,\tau) = \langle T^\tau \chi_A, \chi_B \rangle_\pi / w(A)$
- Probability to stay within A during time τ
  $w(A,A,\tau) = \langle T^\tau \chi_A, \chi_A \rangle_\pi / w(A)$

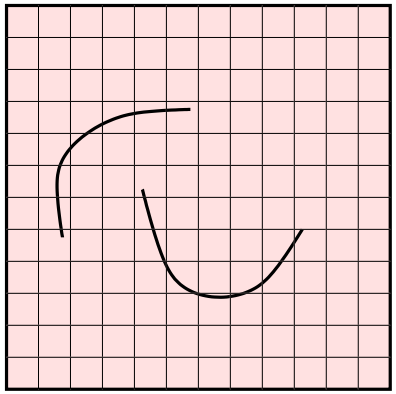
Conformations: almost invariant subsets of the position space Ω

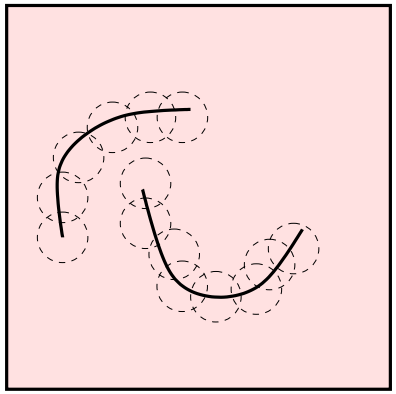
$$w(A,A,\tau) \approx 1 \quad \leftrightarrow \quad T^\tau \chi_A \approx \chi_A$$

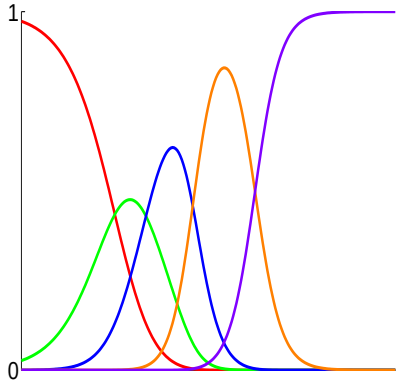
Overview

1. From Molecular Dynamics to Conformation Dynamics
2. Discretization of Position Space
3. Perron Cluster Analysis
4. Error Analysis
5. Softwarecode ZIBgridfree
6. Examples

Discretization







partition of unity

$$\sum_{i=1}^{N} \phi_i(q) = 1, \quad \phi_i(q) \geq 0 \qquad \forall q \in \Omega$$

radial basis functions

$$\phi_i(q) = \frac{\exp(-\alpha\, d(q, q_i)^2)}{\sum_{j=1}^{N} \exp(-\alpha\, d(q, q_j)^2)}, \quad i = 1, \ldots N$$

Eigenvalue Problem

$$T^\tau u = \lambda u, \qquad \lambda \approx 1$$

Galerkin approach:

$$u(q) = \sum_{i=1}^{N} \alpha_i \varphi_i(q)$$

$$\sum_{i=1}^{N} \alpha_i \langle T^\tau \varphi_i, \varphi_j \rangle_\pi = \lambda \sum_{i=1}^{N} \alpha_i \langle \varphi_i, \varphi_j \rangle_\pi, \quad \forall j$$

$$\boxed{\overline{P}\alpha = \overline{S}\alpha\lambda \qquad \text{(symmetric)}}$$

$$\sum_{i=1}^{N} \alpha_i \frac{\langle T^\tau \varphi_i, \varphi_j \rangle_\pi}{\langle \varphi_j \rangle_\pi} = \lambda \sum_{i=1}^{N} \alpha_i \frac{\langle \varphi_i, \varphi_j \rangle_\pi}{\langle \varphi_j \rangle_\pi}, \quad \forall j$$

$$\boxed{P\alpha = S\alpha\lambda \qquad \text{(stochastic)}}$$

Approximation of matrix elements

Approximation of matrices by Monte Carlo importance sampling:

$$\overline{S}(j,i) = \langle \varphi_j, \varphi_i \rangle_\pi = \int_\Omega \varphi_i(q)\varphi_j(q)\pi(q)\,dq \approx \frac{1}{n_j}\sum_{k=1}^{n_j} \varphi_i(q_k^{(j)}),$$

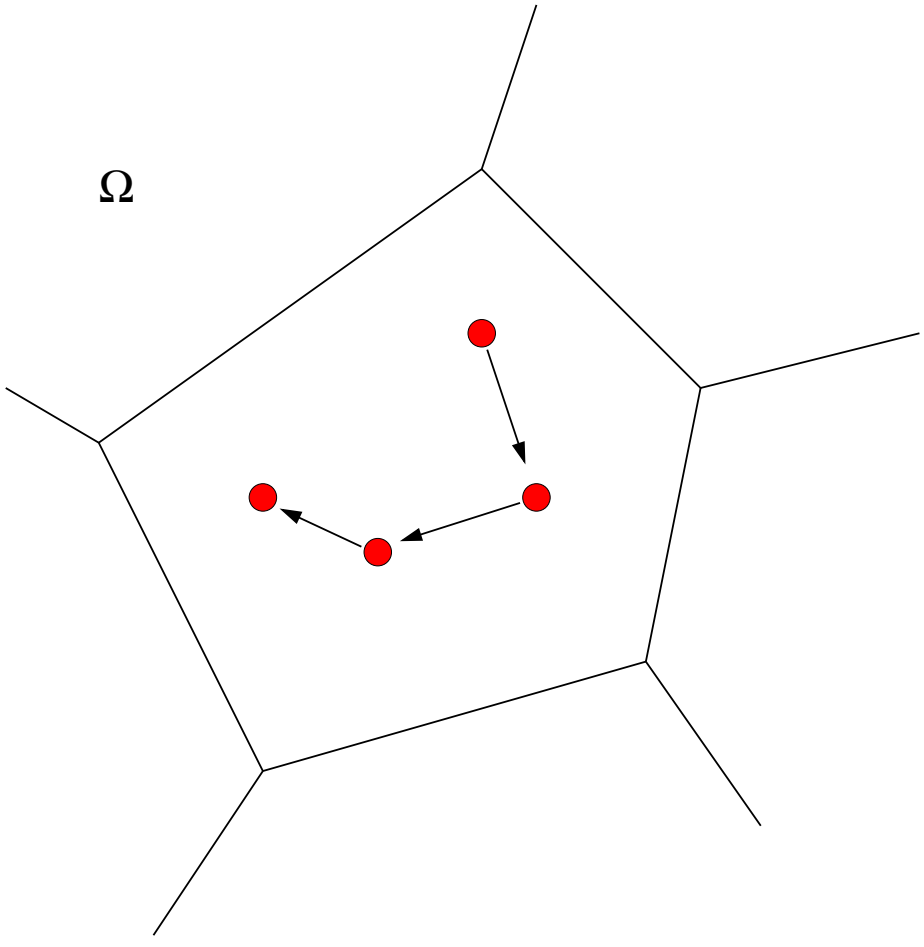
$$\overline{P}(j,i) = \langle T^\tau \varphi_j, \varphi_i \rangle_\pi = \int_\Omega T^\tau\varphi_i(q)\varphi_j(q)\pi(q)\,dq \approx \frac{1}{n_j}\sum_{k=1}^{n_j} \varphi_i(\overline{q}_k^{(j)})$$

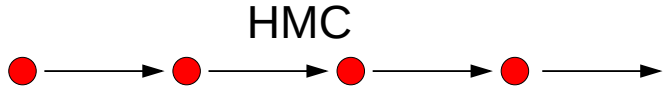
The points $q_k^{(j)}$ are distributed according to partial densities

$$\pi_j(q) \propto \varphi_j(q)\pi(q)$$

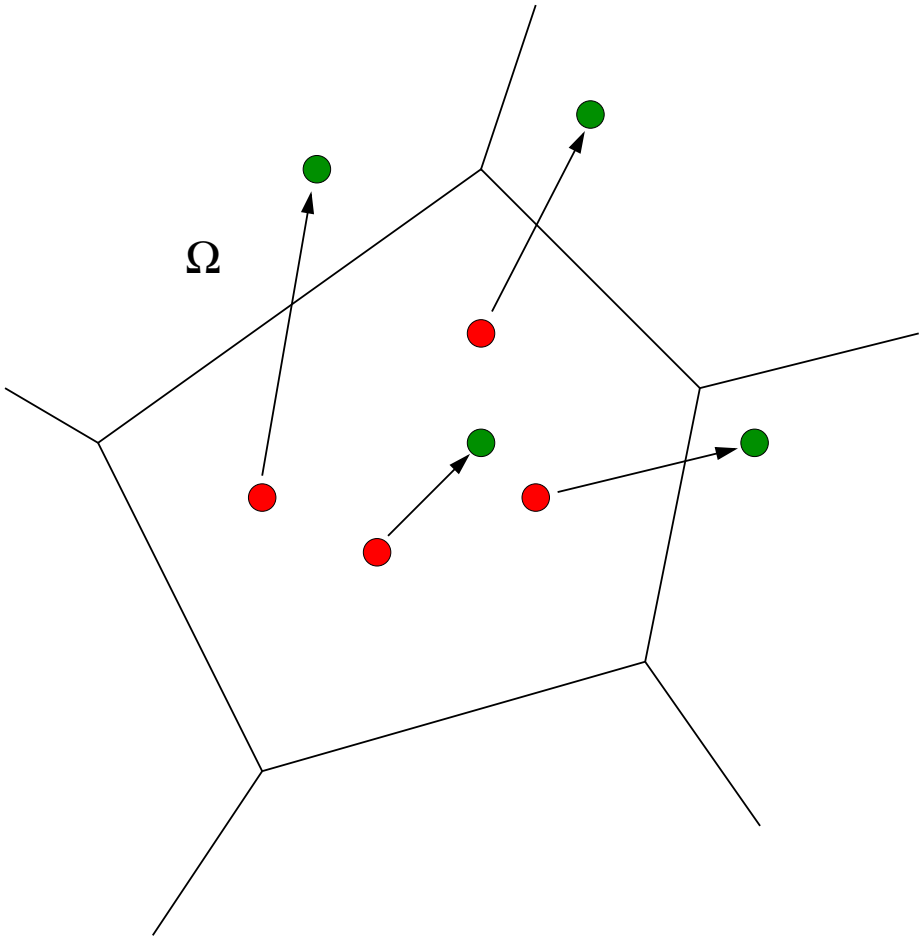
[Weber (2006)]

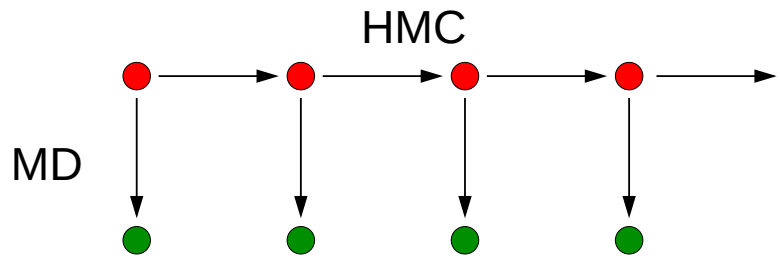
Simulation





Simulation



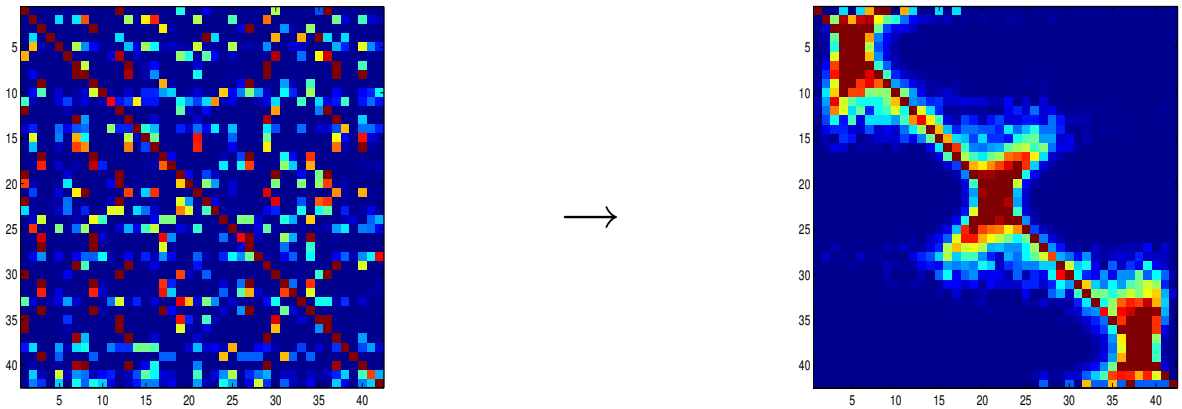


Overview

1. From Molecular Dynamics to Conformation Dynamics
2. Discretization of Position Space
3. Perron Cluster Analysis
4. Error Analysis
5. Softwarecode ZIBgridfree
6. Examples

Crisp Clustering

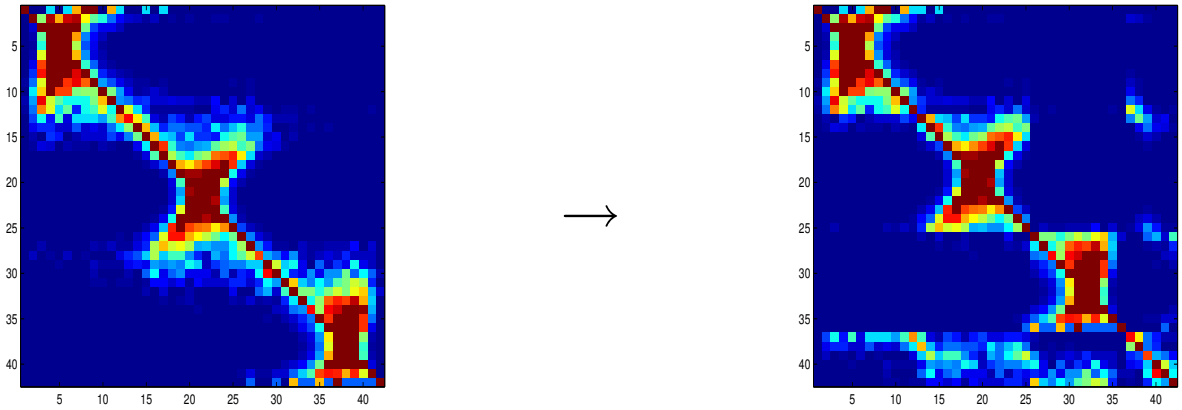
Reordering of states such that there are n_C blocks $\mathcal{C} = \{C_1, \ldots, C_{n_C}\}$.



Grade of membership $\chi(i,j) \in \{0,1\}$:

$$\chi(i,j) = \begin{cases} 1 & \text{if state } i \text{ belongs to cluster } j \\ 0 & \text{else} \end{cases}$$

Soft Clustering



There are states which cannot be assigned uniquely to one of the clusters (transition states!).
Grade of membership $\chi(i,j) \in [0,1]$: state i belongs to cluster j with probability $\chi(i,j)$

- $\sum_{j=1}^{n_C} \chi(i,j) = 1 \quad \forall\, i \in E$
- $0 \leq \chi(i,j) \quad \forall\, i \in E,\, j \in \mathcal{C}$

Robust Perron Cluster Analysis (PCCA+)

MATHEON ZIB

$$PX = X\Lambda, \quad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n_C}), \quad \lambda_1 = 1, \lambda_i \approx 1$$

Transformation of eigenvectors X to membership vectors χ

$$\text{maximize} \quad I(\mathcal{A}) = \sum_{i=1}^{n_C} \frac{\langle \chi_i, \chi_i \rangle_\pi}{\langle \chi_i, e \rangle_\pi} \leq n_C$$

under the conditions

1. $\chi_j(i) \geq 0 \quad \forall i \in \{1, \ldots, N\}, j \in \{1, \ldots, n_C\}$ (positivity),
2. $\sum_{j=1}^{n_C} \chi_j(i) = 1 \quad \forall i \in \{1, \ldots, N\}$ (partition of unity),
3. $\chi = XA$, A regular (invariance)

[Deuflhard, Weber (2005), Röblitz (2008)]

Coarse graining

Transition probabilities between conformations:

$$P_c = (\underbrace{D_c^{-2}\chi^\top D^2 S\chi}_{\mathcal{S}})^{-1}\underbrace{D_c^{-2}\chi^\top D^2 P\chi}_{\mathcal{P}} = (\chi^\top D^2 S\chi)^{-1}\chi^\top D^2 P\chi$$

[Kube, Weber; JCP 2007]

correct propagator, but small negative entries!

objective function in PCCA+:

$$\boxed{I(\mathcal{A}) = \text{trace}(\mathcal{S}) \rightarrow \max}$$

Overview

1. From Molecular Dynamics to Conformation Dynamics
2. Discretization of Position Space
3. Perron Cluster Analysis
4. Error Analysis
5. Softwarecode ZIBgridfree
6. Examples

Perturbation theory

$$\widetilde{P} = P + E, \qquad \widetilde{P}\widetilde{X} = \widetilde{X}\widetilde{\Lambda}$$

Schur decomposition:

$$[X_1, X_2]^H P [X_1, X_2] = \begin{pmatrix} L_1 & H \\ 0 & L_2 \end{pmatrix}, \quad [X_1, X_2]^H E [X_1, X_2] = \begin{pmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix}$$

Perturbation bound:

$$\| \sin \Theta(X_1, \widetilde{X}_1) \| < C \| E_{21} \|$$

[Stewart, Sun (1990)]

Error Estimation

- Row-wise correlated random matrices:

$$\mathbb{E}[E(i,j)] = 0 \quad \text{and} \quad \mathbb{E}[E(i,j)E(k,l)] = \delta_{ik} C_i(j,l)$$

- stochastic norm ([G. W. Stewart, 1990]):

$$\|E\|_s^2 \equiv \mathbb{E}(\|E\|_F^2)$$

$$\|E_{21}\|_s^2 = \sum_{k=1}^{N} \|X_2(k,:)\|_2^2 \text{trace}(X_1^H C_k \overline{X}_1)$$

- probability distribution: $E(i,:) \sim \text{Dir}(\alpha_i)$
- parameter estimation from sampling data (Maximum likelihood estimator)

Adaptive sampling

[Singhal Hinrichs, Pande (2005, 2007)]

Equilibration of sampling effort and uncertainties

$$i = \arg\max_k \operatorname{trace}(X_1^H C_k X_1) \|X_2(k,:)\|_2^2$$

Increase number of sampling points in the selected basis function, update the row, and repeat the analysis.

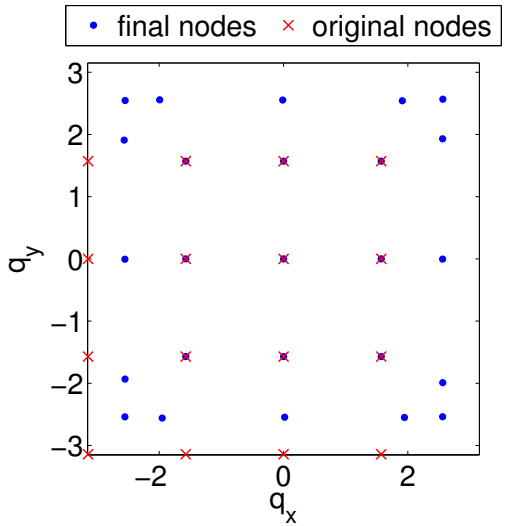
If maximum number of sampling points is reached $\longrightarrow$ refinement

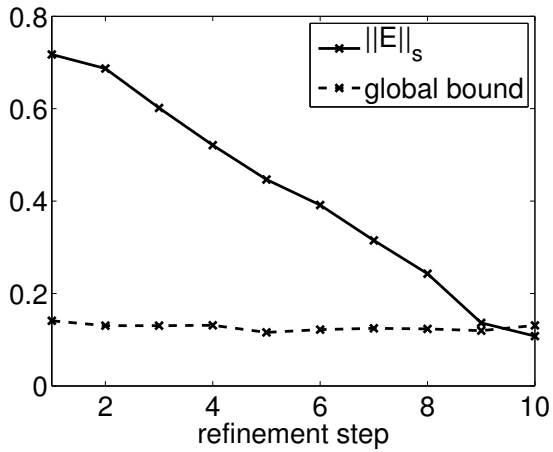
Hierarchical refinement

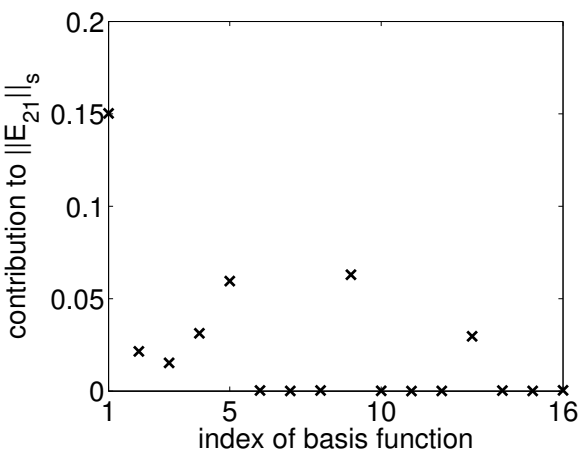
- current basis
$$\{\phi_1, \ldots, \phi_N\} : \Omega \to [0,1]$$
- function selected for refinement: $\phi_k(q)$
- temporal set of basis functions (partition of unity, positivity)
$$\{\widetilde{\phi}_{k1}, \ldots, \widetilde{\phi}_{ks}\} : \Omega \to [0,1]$$
- new basis functions
$$\phi_{ki}(q) := \phi_k(q)\widetilde{\phi}_{ki}(q), \quad i = 1, \ldots, s.$$
$$\{\phi_1, \ldots, \phi_{k-1}, \phi_{k+1}, \ldots, \phi_N, \phi_{k1}, \ldots, \phi_{ks}\}$$

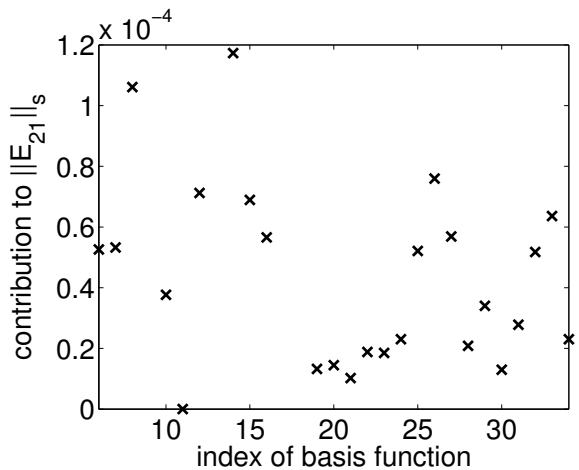
Example

Equilibration of errors and sampling effort









Overview

1. From Molecular Dynamics to Conformation Dynamics
2. Discretization of Position Space
3. Perron Cluster Analysis
4. Error Analysis
5. Softwarecode ZIBgridfree
6. Examples

ZIBgridfree

properties of the software code

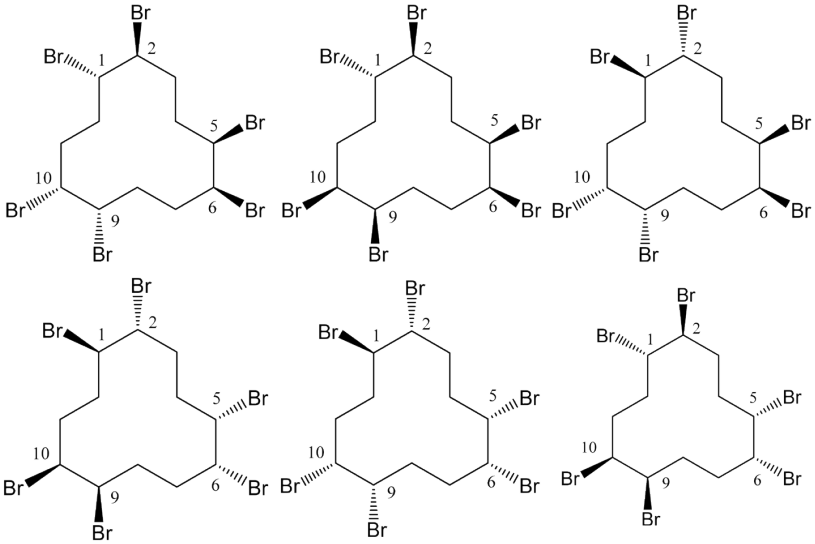
- object oriented program package
- free choice of degrees of freedoms (selection rules)
- restrained modeling (restriction rules)
- different propagators (MD, Smoluchowski, ...)
- mesh-free discretization
- hierarchical refinement
- aggregation/disaggregation (ConfJump)

ZIB spin-off: molConcept GmbH (Marcus Weber)
info@molconcept.com

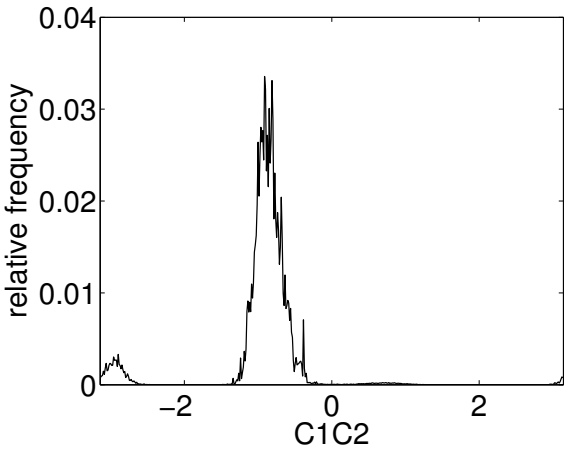
Overview

1. From Molecular Dynamics to Conformation Dynamics
2. Discretization of Position Space
3. Perron Cluster Analysis
4. Error Analysis
5. Softwarecode ZIBgridfree
6. Examples

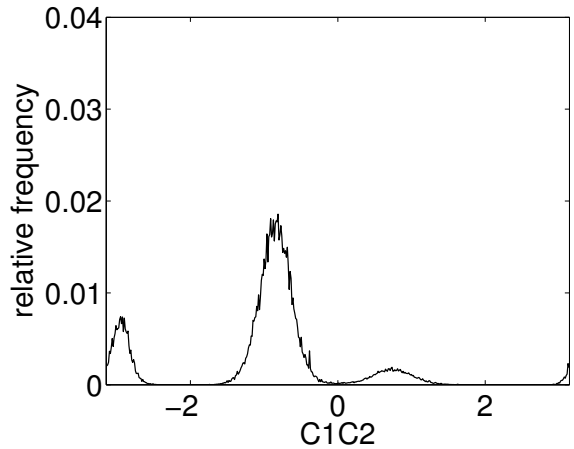
Example: Hexabromocyclododecane



probability for the angle between the bromine atoms to be in gauche or anti position

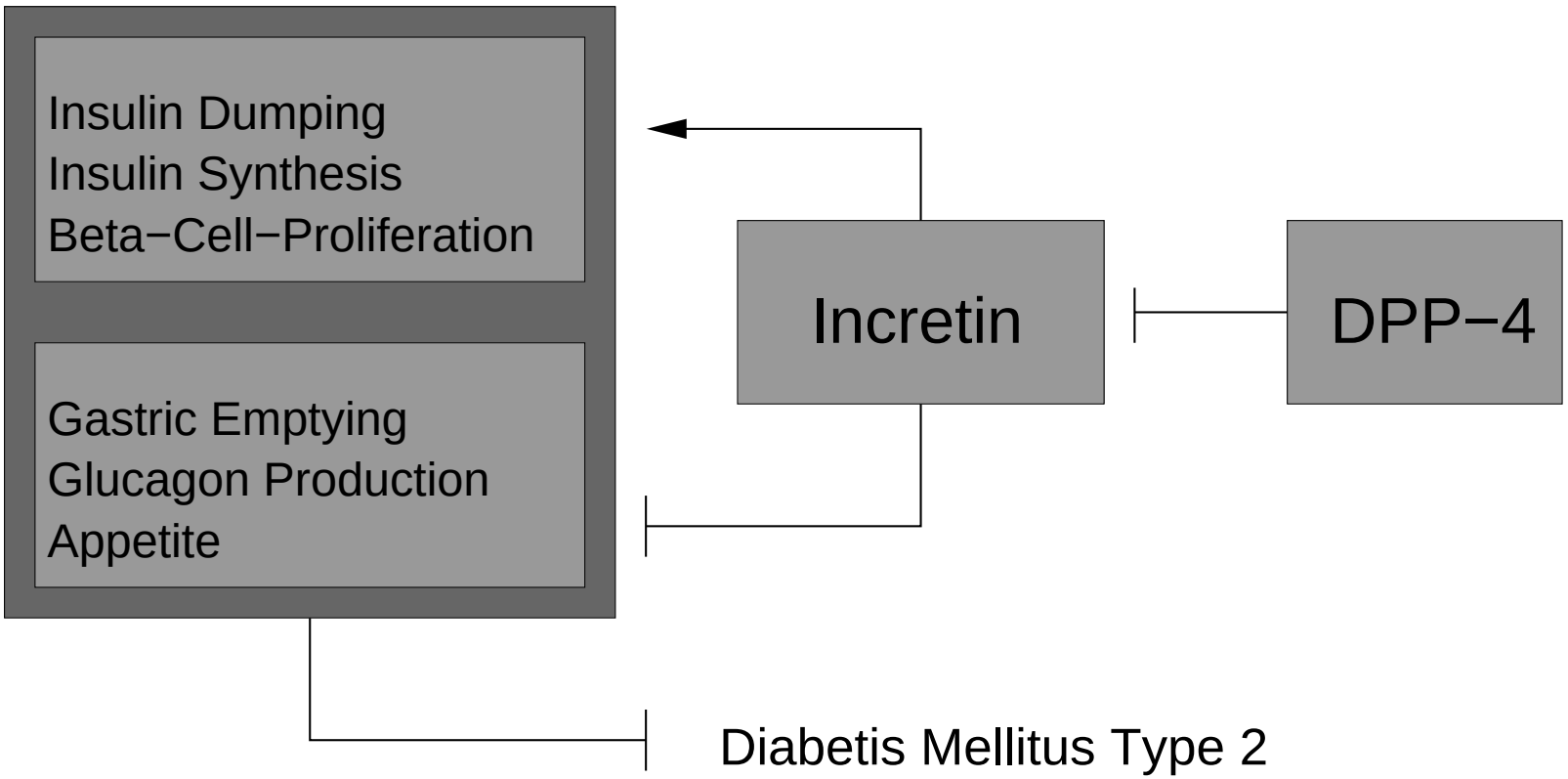


non-adaptive



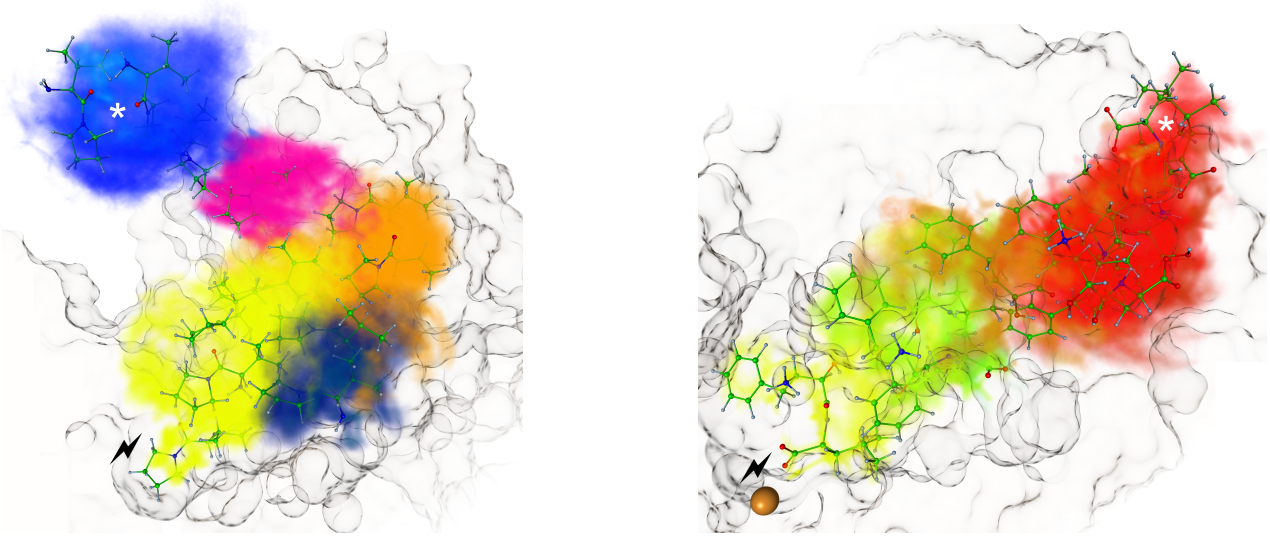
adaptive

DPP-IV Inhibitior



DPP-IV Inhibitior

binding path of ligand Val-Pyr into active center



computing time: 2 days on a laptop!!!

[A. Bujotzek (2008)]

Project Group

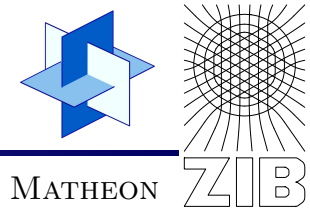
ZIB Computational Drug Design:

Peter Deuflhard	conformation dynamics, PCCA+
Marcus Weber	PCCA+, mesh-free methods
Susanna Röblitz	error analysis, hierarchical refinement, PCCA+
Alexander Bujotzek	DPP-4 inhibition, entropy estimation
Martina Zech	reweighting strategies
Olga Scharkoi	μ-opioid receptor agonists
Vedat Durmaz	toxicology of flame retardants
Karsten Andrae	entropy estimation of PEG derivatives

Cooperations:

DFG Research Center Matheon (Christof Schütte)
SFB-765 "Multivalence" (Rainer Haag)
BAM (Roland Becker)
Charité (Christoph Stein, Hua Fan)

Finally



Thank you for your attention!

Further information

http://www.zib.de/Numerik/DrugDesign/index.en.html